Modeling Systemic Risk with Markov Switching Graphical SUR Models

Daniele Bianchi Monica Billio Roberto Casarin Massimo Guidolin

Journal of Econometrics, forthcoming

Abstract

We propose a Markov Switching Graphical Seemingly Unrelated Regression (MS-GSUR) model to investigate time-varying systemic risk based on a range of multi-factor asset pricing models. Methodologically, we develop a Markov Chain Monte Carlo (MCMC) scheme in which latent states are identified on the basis of a novel weighted eigenvector centrality measure. An empirical application to the constituents of the S&P100 index shows that cross-firm connectivity significantly increased over the period 1999-2003 and during the financial crisis in 2008-2009. Finally, we provide evidence that firm-level centrality does not correlate with market values and it is instead positively linked to realized financial losses.

Keywords: Markov Regime-Switching, Weighted Eigenvector Centrality, Graphical Models, MCMC, Systemic Risk, Network Connectivity
JEL codes: C11, C15, C32, C58

We are grateful to Sylvia Kaufmann (guest editor) and three anonymous referees for their insightful comments and suggestions. Also, we thank Luc Bawuens, Matthias Büechner, Guido Consonni, Fulvio Corsi, Francis X. Diebold, Sylvia Frühwirth-Schnatter, Siam J. Koopman, Fabrizio Lillo, Chris Sims, Mike West, Kamil Yilmaz, and seminar participants at the Warwick Business School, Scuola Normale Superiore of Pisa, HSE of Moscow, the NBER Summer Institute 2015, the NBER-NSF Time Series conference 2015, the 8th Annual SoFiE conference, the SYRTO Conference on Systemic Risk, the 2nd Vienna Workshop on High-Dimensional Time Series in Macroeconomics and Finance, the European Seminar on Bayesian Econometrics ESOBE 2015, and the FMA Orlando 2015, for their helpful comments and advices. Monica Billio and Roberto Casarin acknowledge financial support from the European Union, Seventh Framework Programme FP7/2007-2013 under grant agreement SYRTO-SSH-2012-320270, from the Institut Europlace de Finance under the *Systemic Risk* grant, and from the Italian Ministry of Education, University and Research (MIUR) PRIN 2010-11 grant MISURA. This research used the SCSCF multiprocessor cluster system at University Ca' Foscari of Venice. A previous version of the paper was circulating under the title "Modeling Contagion and Systemic Risk".

Warwick Business School, University of Warwick, Coventry, UK. Daniele.Bianchi@wbs.ac.uk
Department of Economics, University Ca' Foscari of Venice, Venice, Italy. billio@unive.it
Department of Economics, University Ca' Foscari of Venice, Venice, Italy. r.casarin@unive.it
Department of Finance, Baffi-CAREFIN and IGIER, Bocconi University, Milan, Italy. massimo.guidolin@unibocconi.it

1 Introduction

The financial crisis of 2008-2009 has shown that liquidity and valuation shocks may quickly propagate across the economic system and affect financial institutions operating in different markets, with different size and business structure, thus causing widespread losses and domino effects. Understanding the dynamics of cross-asset and cross-equity linkages is therefore of key importance to both systemic risk management purposes and to deal with contagion waves in times of crisis. Systemic risk shocks are conventionally referred to in the network literature as abrupt increases in the density of cross-firm connectivity (see, e.g. Billio, Getmansky, Lo, and Pelizzon 2012, and references therein). Modeling firms' connectedness through network analysis has been recently supported by a series of papers that have shown its in- and out-of-sample superior performance over traditional, correlation-based approaches.[1]

We extend the existing literature in a number of ways. First, we propose a system-wide inferential scheme based on a Markov-switching graphical model that allows us to simultaneously consider all the possible linkages among firms through constraints on the regime-specific conditional dependence structure. Second, we propose an identification scheme for different regimes of cross-firm connectivity based on a novel weighted eigenvector centrality measure, which is related to both the number and the strength of the connections between firms. Third, we provide an asset pricing application based on otherwise standard multi-factor pricing models in which the exposures of the assets to the risk factors (their regression betas) are allowed to change according to the regimes in cross-firm connectivity. This allows us to develop a unified framework where *systematic* and *systemic* risks are not mutually exclusive, in the sense that firm-specific exposures to sources of systematic risk may directly depend on the level of aggregate network connectivity.[2] Finally, we provide a Metropolis-

[1] See, e.g., Forbes and Rigobon (2000), Forbes and Rigobon (2002), Corsetti, Pericoli, and Sbracia (2005), Corsetti, Pericoli, and Sbracia (2011), Billio et al. (2012), Barigozzi and Brownlees (2014), Diebold and Yilmaz (2014),Timmermann, Blake, Tonks, and Rossi (2014), Brownlees, Nualart, and Sun (2014), Diebold and Yilmaz (2015), and Hautsch, Schaumburg, and Schienle (2015) among others.

[2] In this paper, we define systemic risk as the risk caused by the possibility that a firm-specific event may be severe enough to cause pervasive instability in the whole economic system; we define

within-Gibbs sampling scheme which permits to jointly draw both the parameters of the factor pricing model, the latent states and the underlying regime-specific graphs.

Methodologically, we build upon the Gaussian Graphical model for multi-variate systems proposed by Whittaker (1990), Dawid and Lauritzen (1993), Lauritzen (1996), Carvalho, West, et al. (2007), Wang and West (2009), Wang (2010), Rodriguez, Lenkoski, and Dobra (2011), Wang, Reeson, and Carvalho (2011) and Ahelegbey, Billio, and Casarin (2016). In particular, Wang et al. (2011) developed a dynamic matrix-variate graphical model which allows to capture conditional dependencies under time-invariant graphs. We generalize and extend their framework by introducing Markov-switching dynamics in the graph structure within a Seemingly Unrelated Regression (SUR) model. More specifically, we propose a new Markov Switching Graphical SUR (MS-GSUR) which makes it possible to identify different regimes of network connectivity. The regime-switching identification problem is solved by exploiting the graph-theoretic properties of the state-specific conditional dependence structures of the error terms in the model. Thus, we provide a new weighted eigenvector centrality measure, which accounts not only for the number of adjacent nodes, but also for the weights of the edges and for the number of indirect connections between nodes, i.e. the number of walks between nodes. We formally show that our measure can be related to both Bonacich (1972) and communicability (see, e.g., Estrada and Hatano 2008 and Estrada and Hatano 2009), as well as to other measures used in the analysis of topological features of complex networks.

Our empirical application focuses on the cross-section of daily excess returns of the constituents of the S&P100 index over the period 1996-2014. The emphasis on stock returns is motivated by a widespread desire by policy makers and regulators to incorporate the most current information for the purpose of systemic risk measurement: stock prices of largely traded stocks reflect information more rapidly than other non-traded measures such as accounting variables. Informative cross-firm connectivity is estimated on the basis of the residual covariance structure of stock returns,

instead systematic risk as the risk inherent in aggregate market and macroeconomic conditions that cannot be simply diversified away. Moreover, we use the notions of network connectivity, connectedness, and of systemic risk interchangeably.

conditioning on a set of tradable risk factors used in some of the most popular linear asset pricing models, namely the Capital Asset Pricing Model (CAPM), the three-factor model of Fama and French (1993), and Merton's (1973) intertemporal CAPM (I-CAPM). The results are robust across model specifications.

Our main findings reveal that the dynamics of systemic risk can be captured by two regimes, in which a state of high connectedness characterizes the period 1999-2003 (marked by the passing of the Gramm-Leach-Bliley act, the inflating and bursting of the dot.com bubble, and the ensuing financial scandals) and subsequently, the great financial crisis of 2008-2009. We show that a few financial institutions turned out to heavily outweigh other firms in the network during these periods and that shocks to the Financial sector turned out to be the most systemically important. Finally, both a cross-sectional regression and rank-correlation analysis show that market capitalization does not significantly explain the relevance of a given firm within the network. However, firms which are more relevant within the network are more likely to suffer significant losses during periods of high systemic risk.

2 A Markov Switching Graphical SUR Model

Let y_{it} be the stock returns of the ith firm in excess of the risk-free rate at time t, and $\mathbf{x}_{it}$ the m_i-dimensional vector of systematic risk factors. In our baseline formulation, each time series of returns is modeled as a dynamic multi-factor linear model

$$y_{it} = \mathbf{z}_{it}'\boldsymbol{\beta}_i(s_t) + \varepsilon_{it}, \qquad t = 1, \ldots, T, \qquad i = 1, \ldots, n \tag{1}$$

where the matrix $\mathbf{z}_{it} = (1, \mathbf{x}_{it}')'$ includes an intercept plus m_i covariates, $\boldsymbol{\beta}_i(s_t) = (\beta_{i0}(s_t), \beta_{i1}(s_t), \ldots, \beta_{im_i}(s_t))'$ is a $(m_i + 1)$-vector of time-varying regression coefficients, and ε_{it} is an error term that can be identified with firm-specific idiosyncratic risk when $Cov(\mathbf{z}_{it}, \varepsilon_{it}) = \mathbf{0}$. The multi-factor pricing model in (1) is fairly general and represents a reduced-form approximation to a linear pricing kernel (see, e.g., Cochrane 2009). The model in (1) can be re-written in a more compact form as a

SUR, i.e.,

$$\mathbf{y}_t = Z_t' \boldsymbol{\beta}(s_t) + \boldsymbol{\varepsilon}_t, \qquad t = 1, \ldots, T \tag{2}$$

where $\mathbf{y}_t = (y_{1t}, \ldots, y_{nt})'$ is the dependent variable vector, $Z_t = \text{diag}\{\mathbf{z}_{1t}, \ldots, \mathbf{z}_{nt}\}$ is the block-diagonal matrix of covariates, $\boldsymbol{\beta}(s_t)' = (\boldsymbol{\beta}_1'(s_t), \ldots, \boldsymbol{\beta}_n'(s_t))$ is the time-varying coefficient vector of dimension $m = n + \sum_{i=1}^{n} m_i$ and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{nt})'$ is the vector that collects the error terms. We assume that the risk factors are common across stocks and the error terms are independently and normally distributed conditionally on the latent state s_t, $t = 1, \ldots, T$, and have a full time-varying variance-covariance matrix, i.e., $\boldsymbol{\varepsilon}_t \overset{iid}{\sim} \mathcal{N}(\mathbf{0}, \Sigma(s_t))$.

In our model, the time-varying network is identified by the inverse variance-covariance matrix $\Omega(s_t) = \Sigma(s_t)^{-1}$ of the SUR error terms and a set of zero-restrictions implied by an underlying graph $G(s_t) \in \mathcal{G}$, where $\mathcal{G}$ is the space of undirected graphs. More precisely, we introduce a state-dependent graph defined by the ordered pair of disjoint sets $G(s_t) = (V, D(s_t))$ where V is the set of n *vertexes*, or *nodes*, and $D(s_t) \subset V \times V$ defines the set of *edges* in the state s_t.[3] The nodes are state-invariant and represent the firms within the economy and the edges define the presence of interconnections among firms. If two nodes $i, j \in V$ are adjacent in the graph, i.e. $\{i, j\} \in D(s_t)$, then there is an interconnection between two firms. A graph $G(s_t) = (V, D(s_t))$ is complete if $D(s_t) = V \times V$, i.e., it contains all possible edges, that is, all firms are interconnected. Let us define the subgraph G_A of $G(s_t)$ generated by $A \subset V$ as the graph with vertex set A and edge set $K_A \cap D(s_t)$, with $K_A = A \times A$. A *decomposition* of a graph $G(s_t)$ is a partition of V into disjoint nonempty sets (A, B, S) such that the graph generated by S is complete and separates the graphs generated by A and B, that is, any sequence of edges from a vertex in A to a vertex in B contains vertices in S. A sequential decomposition of a graph $G(s_t)$, such that at each iteration i the separator S_i is minimal and the subsets $A_i \subset V$ and $B_i \subset V$ are nonempty, generates the collection $\mathcal{A}(s_t) = \{G_{A_1}, \ldots, G_{A_m}\}$ of m subgraphs that cannot be further decomposed, i.e., its prime components, and the

[3] See, e.g., Bollobás (1998, 2001) for an introduction to graph theory.

collection $\mathcal{B}(s_t) = \{G_{B_1}, \ldots, G_{B_{m'}}\}$ of m' subgraphs generated by the separators. In this paper, we focus on decomposable graphs, where $G(s_t)$ is said to be decomposable either if it is complete or if all its prime components are complete (see, e.g., Dawid and Lauritzen 1993 for further details).

We consider undirected graphs, in which all edges are unordered pairs of vertices. When $G(s_t)$ is an undirected graph, a Gaussian graphical model for the error terms is defined by the assumption that $\boldsymbol{\varepsilon}_t$ is normally distributed with element pairs ε_{it} and ε_{jt}, $i \neq j$ which are conditionally independent given the other error terms and the latent states, i.e.,

$$\varepsilon_{it} \perp \varepsilon_{jt} \,|\varepsilon_{V\setminus\{i,j\}}, s_t \iff \omega_{ij}(s_t) = 0, \quad \forall i \neq j \tag{3}$$

with $\varepsilon_{V\setminus\{i,j\}} = \{\varepsilon_{lt}; l \in V, l \neq i, j\}$ and $\omega_{ij}(s_t)$ $(i,\, j = 1, \ldots, n)$ the $(i,\, j)$-th element of the precision matrix $\Omega(s_t)$. Hence, $\Omega(s_t) \in \mathcal{M}(G(s_t))$ with $\mathcal{M}(G(s_t))$ the set of all positive-definite symmetric matrices with elements equal to zero for all $\{i, j\} \notin D(s_t)$.

The Gaussian nature of our framework might present limitations as it does not permit to capture co-dependencies in higher moments. While a distinctive attention to covariances and correlations is typical of a classical finance (e.g., of the mean-variance paradigm) approach, clearly this overlooks other important aspects. However, given our goal to use standard I-CAPM-style models to differentiate the notions of systematic risk from systemic risk, our choice seems a sensible one as a first step of analysis. The time-variation in the parameters of the MS-GSUR model is driven by a latent first-order Markov chain process s_t with time-homogeneous transition probabilities $\text{Prob}(s_t = j|s_{t-1} = i) = \pi_{ij}$, $i, j = 1, \ldots, K$, for all $t = 1, \ldots, T$. The choice of a regime-dependent dynamics is motivated by its popularity in the finance literature as it allows for an intuitive economic interpretation of different market phases (see, e.g., Guidolin and Timmermann 2008). The Markov switching dynamics of the SUR coefficients can be represented as

$$\boldsymbol{\beta}(s_t) = \sum_{k=1}^{K} \boldsymbol{\beta}_k \mathbb{I}_{\{k\}}(s_t), \tag{4}$$

where $\mathbb{I}_{\{k\}}(s_t)$ is an indicator function that takes a value of one when the state s_t takes value k at time t, and zero otherwise, and $\boldsymbol{\beta}_k$ is the vector of regime-specific betas. We assume that for each state $s_t = k$ there is a regime-specific covariance matrix $\Sigma(s_t)$ constrained by the state-specific graph $G_k = (V, D_k)$, such that:

$$\Sigma(s_t) = \sum_{k=1}^{K} \Sigma_k \mathbb{I}_{\{k\}}(s_t), \tag{5}$$

The regime-dependent feature of the covariance structure allows us to address potential heteroskedasticity biases that are typical of correlation-based measures, as discussed in Forbes and Rigobon (2002). Also, the topological features of the state-specific graph G_k play a crucial role in the estimation of the model since they allow us to identify regimes of low vs. high systemic risk. The factor model specification in (2)-(5) implies that systematic and systemic risks are not mutually exclusive. In particular, while the exposures to systematic risks, i.e., the betas, depend on the latent state, the latter, as shown below, is directly identified from the network defined by the residuals, which on their turn depend on the regression betas. In this respect, the relationship between systematic and systemic risk captured by the model is contemporaneous and not causal, and posterior inference is interpretable only as a result of an interplay between systematic and systemic risks, which provides evidence about the most likely combination of systematic and systemic risk.

3 Inference on Networks and Parameters

Let G_k, $k = 1, \ldots, K$ be a sequence of decomposable graphs, $\mathcal{A}_k = \{A_{1,k}, \ldots, A_{n_{A,k},k}\}$ the set of $n_{A,k}$ complete prime components of G_k and $\mathcal{B}_k = \{B_{1,k}, \ldots, B_{n_{B,k},k}\}$ the set of $n_{B,k}$ separators of G_k. If the joint distribution of excess stock returns is Markov with respect to a decomposable graph G_k, the joint density of $\mathbf{y}_t$ given $s_t = k$ factorizes as

$$p(\mathbf{y}_t | Z_t, \boldsymbol{\beta}_k, \Sigma_k, G_k, s_t) = \prod_{g \in \mathcal{A}_k} p\left(\mathbf{y}_{gt} | Z_{gt}, \boldsymbol{\beta}_{gk}, \Sigma_{gk}, s_t\right) / \prod_{g \in \mathcal{B}_k} p\left(\mathbf{y}_{gt} | Z_{gt}, \boldsymbol{\beta}_{gk}, \Sigma_{bk}, s_t\right) \tag{6}$$

For each subgraph $g \in \mathcal{A}_k$ or $g \in \mathcal{B}_k$, $\mathbf{y}_{gt} = (y_{it} : i \in g)$, Z_{gt} is the corresponding matrix of covariates, $\boldsymbol{\beta}_{gk} = (\beta_{ik} : i \in g)$ the subset of factor loadings, and Σ_{gk} the relative block of the covariance matrix of the residuals from Σ_k. Each term in (6) has a multivariate normal distribution, $\mathbf{y}_g \sim \mathcal{N}(Z_g'\boldsymbol{\beta}_{gk}, \Sigma_{gk})$, where $\Omega_{gk} = \Sigma_{gk}^{-1}$ is a full positive-definite symmetric matrix. Given the graph G_k, the joint distribution (6) is completely defined by the component-specific marginal betas, covariates, and covariance matrices (see, e.g., Giudici, Green, and PJ 1999 and Carvalho et al. 2007).

3.1 Prior Specification

Dawid and Lauritzen (1993) define a class of probability distributions for covariance matrices on decomposable graphs called hyper-inverse Wishart. Conditional on the graph G_k, the prior distribution of the k-th state covariance matrix Σ_k is the hyper-inverse Wishart

$$\Sigma_k \sim \mathcal{HIW}_{G_k}(c_k, C_k), \tag{7}$$

with c_k and C_k the degrees of freedom and the location parameters, respectively. We denote this prior distribution with $p(\Sigma_k|G_k)$. For each state $s_t = k$, the hyper-inverse Wishart represents the unique conjugate "local prior" for complete prime components that are inverse Wishart distributed (see Carvalho, West, and Massam 2007 for a detailed discussion). By generating the tree representation of the prime components, the density of the hyper-inverse Wishart for Σ_k conditional on G_k has the following expression

$$p(\Sigma_k|G_k) = \prod_{j=1}^{n_{A,k}} p(\Sigma_{A_{j,k}}) \prod_{i=1}^{n_{B,k}} p(\Sigma_{B_{i,k}})^{-1} \tag{8}$$

where for each prime component $\Sigma_{A_{j,k}} \sim \mathcal{IW}\left(c_k, C_{A_{j,k}}\right)$, so that the density is given by

$$p(\Sigma_{A_{j,k}}) \propto |\Sigma_{A_{j,k}}|^{-\left(c_k+2T_{A_{j,k}}\right)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma_{A_{j,k}}^{-1} C_{A_{j,k}})\right\} \tag{9}$$

where $T_{A_{j,k}} = Card(A_{j,k})$ is the cardinality of $A_{j,k}$, and $C_{A_{j,k}}$ is the jth diagonal block of C_k corresponding to $\Sigma_{A_{j,k}}$ (see Hammersley and Clifford 1971 and Dempster 1972). The initial sparse-inducing prior over the graph structure is defined as a Bernoulli distribution with parameter ψ on each edge inclusion probability. That is, a n-node graph $G_k = (V, D_k)$ has a prior probability

$$p(G_k) \propto \prod_{i,j} \psi^{d_{ij,k}} (1-\psi)^{(1-d_{ij,k})} \propto \psi^{Card(D_k)} (1-\psi)^{N-Card(D_k)} \tag{10}$$

where $d_{ij,k} = 1$ if $\{i, j\} \in D_k$ for $i \neq j$, and 0 otherwise, and $Card(D_k) = \sum_{i,j \in V} d_{ij}/2$ denotes the cardinality of D_k, which is the number of edges (see Bollobás 1998, Chapter 1). Given that the graph is undirected, the cardinality of D_k is given by $\sum_{i,j \in V} d_{ij}/2$, which is equal to $N = n(n-1)/2$ if the n-node graph is complete, i.e., there is an edge between all pairs of nodes. The prior (10) has its peak at $N\psi$ providing a flexible way to directly control for prior model complexity.[4] To induce sparsity and hence obtain a parsimonious representation of the interdependence structure implied by a graph, we choose $\psi = 2/(n-1)$ which provides a prior mode in correspondence to n edges. In addition to this baseline specification, we experiment with two alternative prior specifications for $p(G_k)$ which imply either an empty graph, or a complete graph, i.e., $d_{ij,k} = 1$ for $i,j = 1, \ldots, n$, $i \neq j$. The posterior distribution of the set G_k tends to reach a similar posterior median for different starting priors. The prior for the factor betas is chosen to be independent of the covariance structure:

$$\boldsymbol{\beta}_k \sim \mathcal{N}(\mathbf{m}_k, M_k) \tag{11}$$

We choose prior distributions for the regression parameters centered around zero to be rather uninformative and common across states. The prior distribution for the kth row of the transition matrix Π, i.e., $\boldsymbol{\pi}_k = (\pi_{k1}, \ \ldots, \ \pi_{kK})'$ is a Dirichlet, i.e.,

$$\boldsymbol{\pi}_k \sim \mathcal{D}ir(c_{k1}, \ldots, c_{kK}) \tag{12}$$

[4]Alternatively, a uniform prior might have been used instead. However, as pointed out in Jones, Carvalho, Dobra, Hans, Carter, and West (2005), a uniform prior over the space of all graphs is biased towards a graph with half of the total number of possible edges.

with c_{lk} the concentration hyper-parameter.

3.2 Posterior Approximation

Let us denote with $\mathbf{y}_{\tau:t} = (\mathbf{y}_\tau, \ldots, \mathbf{y}_t)$, $\tau \leq t$, the sequence of observations between τ and t, with $G = (G_1, \ldots, G_K)$ and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_K)$ the collections of state-specific graphs and parameters, respectively, where $\boldsymbol{\theta}_k = (\boldsymbol{\beta}_k, \Sigma_k, \boldsymbol{\pi}_k)$, $k = 1, \ldots, K$. The complete likelihood is then defined as

$$p\left(\mathbf{y}_{1:T}, \mathbf{s}_{1:T} | \boldsymbol{\theta}, G\right) = \prod_{t=1}^{T} p\left(\mathbf{y}_t | s_t, \boldsymbol{\theta}, G\right) p\left(s_t | s_{t-1}, \boldsymbol{\theta}\right) p\left(s_0\right) \tag{13}$$

With reference to our application, the marginal likelihood accommodates the existence of fat tails in the distribution of excess stock returns. Indeed, given the local conjugate priors and the hyper-Markov structure of the graph G_k, one can show that the marginal distribution conditional on the sequence of states is an *hyper Student-t* (see, e.g., Dawid and Lauritzen 1993 for details).

Let $p(\boldsymbol{\theta}, G) \propto \prod_{k=1}^{K} p(\boldsymbol{\beta}_k) p(\Sigma_k | G_k) p(G_k) p\left(\boldsymbol{\pi}_k\right)$ be the joint prior distribution; then the joint posterior is $p(\boldsymbol{\theta}, G | \mathbf{s}_{1:T}, \mathbf{y}_{1:T}) \propto p\left(\mathbf{y}_{1:T}, \mathbf{s}_{1:T} | \boldsymbol{\theta}, G\right) p(\boldsymbol{\theta}, G)$. Because such distribution is not tractable, the estimator of the parameters and graphs cannot be obtained in analytical form. We approximate the posterior distribution by implementing a multi-move Gibbs sampling algorithm (see, e.g., Roberts and Sahu 1997 and Casella and Robert 2004), where the graph structure, the hidden states, and the parameters are all sampled in blocks. At each iteration, the Gibbs sampler sequentially cycles through the following steps:

1. sample $\mathbf{s}_{1:T}$ given the graphs G, the parameters $\boldsymbol{\theta}$ and the observations $\mathbf{y}_{1:T}$;
2. sample Σ_k given $\mathbf{s}_{1:T}$, $\boldsymbol{\beta}_k$, G_k and $\mathbf{y}_{1:T}$, for $k = 1, \ldots, K$,
3. sample G_k given Σ_k, $\mathbf{s}_{1:T}$ and $\mathbf{y}_{1:T}$, for $k = 1, \ldots, K$,
4. sample $\boldsymbol{\beta}_k$ given $\mathbf{s}_{1:T}$, Σ_k, G_k and $\mathbf{y}_{1:T}$, for $k = 1, \ldots, K$,
5. sample Π given $\mathbf{s}_{1:T}$.

We extract the latent states s_t, $t = 1, \ldots, T$ by using a forward-filtering backward-

sampling (FFBS) algorithm (see Frühwirth-Schnatter 1994 and Carter and Kohn 1994). Because the latent state is discretely valued, the FFBS is applied in its Hamilton (1994)'s form.

Denote with $\mathcal{T}_k = \{t : s_t = k\}$ the index set for observations in regime k and with $T_k = Card(\mathcal{T}_k)$ its cardinality. Also, let $\mathbf{e}_{tk} = \mathbf{y}_t - Z_t'\boldsymbol{\beta}_k$ be the residuals conditional on the state $s_t = k$ and $E_k^* = \sum_{t\in\mathcal{T}_k} \mathbf{e}_{tk}\mathbf{e}_{tk}'$. Given the local conjugate prior in (7), the posterior for Σ_k is

$$p\left(\Sigma_k|\mathbf{y}_{1:T}, \boldsymbol{\theta}, \mathbf{s}_{1:T}, \boldsymbol{\beta}_k\right) \propto \mathcal{HIW}_{G_k}\left(c_k + T_k, C_k + E_k^*\right), \tag{14}$$

see Appendix A for a proof. To sample the graphs $G_k, k = 1, \ldots, K$, we compute the unnormalized posterior over the graphs $p(G_k|\mathbf{y}_{1:T}, \mathbf{s}_{1:T}) \propto p(\mathbf{y}_{1:T}|\mathbf{s}_{1:T}, G_k)p(G_k)$, for any specified state k (see, e.g. Giudici et al. 1999 and Jones et al. 2005). As in the proof of (14), given the prior independence assumption of the parameters across regimes, one can show that

$$p\left(G_k|\mathbf{y}_{1:T}, \mathbf{s}_{1:T}\right) \propto p\left(G_k\right) \int\int p(\mathbf{y}_{\mathcal{T}_k}|\mathbf{s}_{\mathcal{T}_k}, \boldsymbol{\beta}_k, \Sigma_k)p(\boldsymbol{\beta}_k)p(\Sigma_k|G_k)d\boldsymbol{\beta}_k d\Sigma_k \tag{15}$$

where

$$p(\mathbf{y}_{\mathcal{T}_k}|\mathbf{s}_{\mathcal{T}_k}, \boldsymbol{\beta}_k, \Sigma_k) \propto |\Sigma_k|^{-T_k/2} \exp\left\{-\frac{1}{2}\text{tr}\left(\Sigma_k^{-1}E_k^*\right)\right\}.$$

To evaluate this integral, we follow Chib (1995) and Wang (2010) and approximate the marginal likelihood via a local-move Metropolis-Hastings step based on the conditional posterior $p(G_k|\mathbf{y}_{1:T}, \mathbf{s}_{1:T})$. A candidate G_k' is sampled from a proposal distribution $q(G_k'|G_k)$ and accepted with probability

$$\alpha = \min\left\{1, \frac{p(G_k'|\mathbf{y}_{1:T}, \mathbf{s}_{1:T})q(G_k|G_k')}{p(G_k|\mathbf{y}_{1:T}, \mathbf{s}_{1:T})q(G_k'|G_k)}\right\} \tag{16}$$

where $p(G_k|\mathbf{y}_{1:T}, \mathbf{s}_{1:T})$ has been defined in (15). The add/delete edge move proposal is rather accurate, although such accuracy comes at the price of a substantial computational burden. The full conditional posterior distribution of the state-specific SUR

coefficient $\boldsymbol{\beta}_k$ is conjugate and defined as

$$p\left(\boldsymbol{\beta}_k|\Sigma_k, \mathbf{y}_{1:T}, \mathbf{s}_{1:T}\right) \propto \mathcal{N}\left(M_k^*\left(\sum_{t\in\mathcal{T}_k} Z_t\Omega_k\mathbf{y}_t + M_k^{-1}\mathbf{m}_k\right), M_k^*\right) \tag{17}$$

with $M_k^* = \left(\sum_{t\in\mathcal{T}_k} Z_t\Omega_k Z_t' + M_k^{-1}\right)^{-1}$, and $\Omega_k = \Sigma_k^{-1}$ is the inverse covariance matrix given the underlying graph structure, G_k. Finally, the conjugate Dirichlet prior for the rows of the transition probability $\boldsymbol{\pi}_k = (\pi_{k1}, \ldots, \pi_{kK})'$ is updated as follows

$$p(\boldsymbol{\pi}_k|\mathbf{s}_{1:T}) \propto \mathcal{D}ir\left(c_{k1} + N_{k1}, \ldots, c_{kK} + N_{kK}\right) \tag{18}$$

where N_{lk}, $k = 1, \ldots, K$, is the empirical distribution which counts the transition between the lth and the kth latent discrete states, i.e., $N_{lk} = \sum_{t=1}^{T} \xi_{lk,t}$ with $\xi_{lk,t} = \mathbb{I}_{\{k\}}(s_t)\,\mathbb{I}_{\{l\}}(s_{t-1})$.

3.3 States Identification via Eigenvector Centrality

The likelihood function in (13) remains unchanged with respect to any state permutation. Therefore, under an invariant prior specification, the posterior will also be invariant to any state permutation (see Frühwirth-Schnatter 2006 for a review of label-switching and identification issues). We propose a state-identifying restriction based on a weighted eigenvector centrality measure. In its general form, the eigenvector centrality of the ith firm/stock, is a quantity proportional to the sum of the centralities of the neighbors of a vertex, so that a vertex may display a high centrality either by being connected to a lot of other vertexes or by being connected to other vertexes that are themselves highly central.

As in many other fields (e.g., biology, neuroscience, and operations research), where complex networks have been studied, it is possible to assign to each edge of the graph a weight proportional to the intensity of the connections among the various elements of the network. Appropriate metrics combining weighted and topological observables have been discussed in the literature (see Rubinov and Sporns 2010 for a review); as argued in Barrat, Barthlemy, Pastor-Satorras, and Vespignani (2004),

network metrics based on weighted edges allow us to provide a better description of the hierarchies and organizational principles at the basis of the architecture of the networks. In the Gaussian graphical model literature, covariances, precisions, and the topological features of the graph, such as the paths between pairs of nodes, see, e.g., Jones and West (2005), have been routinely combined together with the goal of assessing the centrality of each node.

In our financial application, the information about the existence of a financial linkage between pairs of stocks is encoded in the presence or absence of an edge between two nodes, while the strength of the linkage is measured by the covariance between pairs of stocks. Thus, we construct a state-specific weighted graph defined as $\tilde{G}_k = (V, D_k, \tilde{\Sigma}_k)$, where $\tilde{\Sigma}_k$ is a real-valued symmetric matrix in which each entry is

$$\tilde{\sigma}_{ij,k} = \begin{cases} \sigma_{ij,k} & \text{if } (i,j) \in D_k \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

that is the weight $\tilde{\sigma}_{ij,k}$ assigned to each pair of nodes $\{i,j\} \in V \times V$ is equal to the corresponding covariance if they are connected in state $k \in K$. The relative centrality score for the ith firm in the weighted graph $\tilde{G}_k$, i.e., $\gamma_{i,k}$, can be defined as

$$\lambda_k \gamma_{i,k} = \sum_{j=1}^{n} \tilde{\sigma}_{ij,k} \gamma_{j,k}, \tag{20}$$

where λ_k is some constant. With a small re-arrangement, this measure can be re-written in a more compact form as the eigenvector equation $\tilde{\Sigma}_k \boldsymbol{\gamma}_k = \lambda_k \boldsymbol{\gamma}_k$ where $\boldsymbol{\gamma}_k = (\gamma_{1,k}, \ldots, \gamma_{n,k})'$. Since $\tilde{\Sigma}_k$ is real and symmetric, a unique solution is guaranteed to exist by the Perron-Frobenius theorem.[5] Our average weighted eigenvector centrality of the graph $\tilde{G}_k$ is defined as

$$\overline{q}(\tilde{G}_k) = \frac{1}{n} \sum_{i=1}^{n} \gamma_{in,k} \tag{21}$$

[5] See, e.g., Bollobás (1998), Chapter 8, Theorem 5, for a description of the relationship between the maximal eigenvalue and the minimal and maximal degree of a graph.

where $\gamma_{in,k}$ is the i-th element of the eigenvector $\boldsymbol{\gamma}_{n,k}$ corresponding to the largest eigenvalue $\lambda_{n,k}$. The advantage of (21) is that it accounts not only for the number of connections of each node with the adjacent nodes, but also for its weight and for the weights of the indirect connections with other nodes on the graph. The notion of indirect connection between nodes is clarified by the following definitions of walks and paths between nodes (see Bollobás 1998, Chapter. 1, for further details):

Definition 1. *A walk* $p = (i_0, e_1, \ldots, e_l, i_l)$*, between two vertices* i, j *of* G_k *is identified by an alternating sequence of (not necessary different) vertices* $V(p) = \{i_0, i_1, \ldots, i_l\}$*, with* $i_0 = i$ *and* $i_l = j$*, and edges* $D_k(p_{i_0 i_l}) = \{e_1, \ldots, e_l\} \subset D_k$*, with* $e_1 = (i_0, i_1)$ *and* $e_l = \{i_{l-1}, i_l\}$*. The number of edges* $Card(D_k(p_{ij})) = l$ *in a walk is called "walk length". A walk of length* l *is called* l*-walk.*

Definition 2. *A path* p_{ij} *between vertices* i, j *of* G_k *is a walk with distinct elements in its vertex set. The shortest-path* p^*_{ij} *between two vertices* i, j *is* $\min_l\{p = (i_0, e_1, \ldots, e_l, i_l), l \geq 1\}$*, that is, the path with minimum length.*

Proposition 1 provides an interpretation for our measure by showing that it can be made arbitrarily close to a weighted sum over walks, where weights are inversely related to the length of each walk.

Proposition 1. *Let* $\tilde{G}_k = (V, D_k, \tilde{\Sigma}_k)$ *be a weighted undirected graph with vertex set* $V = \{1, \ldots, n\}$*, edge set* D_k *and real-valued weight matrix* $\tilde{\Sigma}_k$*. Denote with* $\lambda_{1,k} \leq \lambda_{2,k} \leq \ldots \leq \lambda_{n,k}$ *the eigenvalues of* $\tilde{\Sigma}_k$*, with* $\boldsymbol{\gamma}_{j,k}$*,* $j = 1, \ldots, n$ *the associated eigenvectors, and with* $\overline{q}(\tilde{G}_k)$ *be average eigenvector centrality. Assume the maximal eigenvalue* $\lambda_{n,k}$ *has multiplicity one, and define* $I_\delta(1/\lambda_{n,k}) = \{b : 0 < 1/\lambda_{n,k} - b < \delta\}$*,* $\delta > 0$ *as the neighborhood of* $1/\lambda_{n,k}$*. Then* $\forall \varepsilon > 0$*,* $\exists \delta > 0$ *such that for* $\forall b \in I_\delta(1/\lambda_{n,k})$

$$\left| \frac{1}{n\kappa(b,k)} \sum_{l=1}^{\infty} b^{l-1} \tilde{\sigma}_{l,k} - \overline{q}(\tilde{G}_k) \right| < \varepsilon \tag{22}$$

where

$$\tilde{\sigma}_{l,k} = \sum_{i,j \in V} \sum_{p \in P_{ij,k}^{(l)}} \tilde{\sigma}(p)$$

is the sum of the walk weights over the set $P_{ij,k}^{(l)}$ *of all possible* l*-walks between vertices* i *and* j *for a given state* $k = 1, \ldots, K$*, where* $\tilde{\sigma}(p) = \prod_{r=1}^{l} \tilde{\sigma}_{i_{r-1} i_r, k}$ *is the weight product over the edges of a walk* $p = (i_0, e_1, \ldots, e_l, i_l)$ *and*

$$\kappa(b, k) = \sum_{i=1}^{n} \frac{b\lambda_{i,k}}{1 - b\lambda_{i,k}}$$

is a normalizing factor.

Proof See Appendix A.

Other centrality measures have been proposed in the literature, such as average node degree, closeness, and betweenness. However, these measures may fail in identifying alternative regimes of systemic risk. Degree centrality gives a simple count of the number of connections any given firm or asset maintains, without effectively discriminating the relative importance of these connections with respect to the whole network. However, linkages across firms are arguably not all alike. Analogously, closeness and betweenness are measures of centrality based on the shortest paths between one node and all other nodes in the graph. These measures implicitly assume simplistic and pre-determined paths and may be severely inappropriate for the analysis of the transition mechanism of economic shocks which in general have feedback effects and are unlikely to be restricted to follow specific paths. Our centrality measure addresses these issues and takes into account that an edge or the shortest path between two nodes does not measure the full level of connectedness between two nodes.

In addition, our measure shares some features with the communicability measures for complex networks proposed in Estrada and Hatano (2008, 2009) and more generally with other global connectivity measures (e.g., see Han, Escolano, Hancock, and Wilson 2012, Qi, Fuller, Wu, Wu, and Zhang 2012, and Qi, Fuller, Luo, and Zhang 2015) by admitting decompositions in walks, spanning trees, and circuits. More

specifically, the following decomposition holds:

Corollary 1. *Let* $\tilde{\sigma}(p) = \prod_{r=1}^{l} \tilde{\sigma}_{i_{r-1}i_r,k}$ *be the weight product over the edges of a walk* $p = (i_0, e_1, \ldots, e_l, i_l)$*, and* $\overline{q}(\tilde{G}_k)$ *the weighted eigenvector centrality measure of the graph* $\tilde{G}_k$*. Define* $I_\delta(1/\lambda_{n,k}) = \{b : 0 < 1/\lambda_{n,k} - b < \delta\}$*,* $\delta > 0$*, the neighborhood of the inverse maximal eigenvalue* $1/\lambda_{n,k}$*. Then* $\forall \varepsilon > 0$*,* $\exists \delta > 0$ *such that for* $\forall b \in I_\delta(1/\lambda_{n,k})$

$$\left| \frac{1}{n\kappa(b,k)} \sum_{i,j \in V} \left(\sum_{p \in P^*_{ij,k}} b^{l^*_{ij}-1} \tilde{\sigma}(p) + \sum_{l > l^*_{ij}} b^{l-1} \sum_{p \in P^{(l)}_{ij,k}} \tilde{\sigma}(p) \right) - \overline{q}(\tilde{G}_k) \right| < \varepsilon \qquad (23)$$

where $P^*_{ij,k}$ *and* l^*_{ij} *denote the set of the shortest paths between nodes* i *and* j *of the graph* $\tilde{G}_k$ *and their length, respectively.*

Proof See Appendix A.

The first term in (23) reflects the connectivity due to the shortest paths and degree distribution, while the second term reflects the connectivity or influence between nodes at a global level, reflecting losses spreading into the financial system forward and backward several times from a source to a destination. Thus, the measure we propose provides a better representation of more complex structures such as scale-free or small-world networks. This aspect is also reflected by a different decomposition of our measure into intra-cluster and inter-cluster communicability terms. The following decomposition shows that our measure also naturally accounts for community structures (see Fortunato 2010 for a graph-theoretic definition of community).

Corollary 2. *Let* $\overline{q}(\tilde{G}_k)$ *be the weighted eigenvector centrality measure of the graph* $\tilde{G}_k$*. Define* $I_\delta(1/\lambda_{n,k}) = \{b : 0 < 1/\lambda_{n,k} - b < \delta\}$*,* $\delta > 0$*, as the neighborhood of the inverse maximal eigenvalue* $1/\lambda_{n,k}$*. Then* $\forall \varepsilon > 0$*,* $\exists \delta > 0$ *such that for* $\forall b \in I_\delta(1/\lambda_{n,k})$

$$\left| \frac{1}{n\kappa(b,k)} \sum_{i \in V} \frac{\lambda_{i,k}}{1 - b\lambda_{i,k}} \left(\varphi^{(1)}_{i,k} - 2\varphi^{(2)}_{i,k} \right) - \overline{q}(\tilde{G}_k) \right| < \varepsilon \qquad (24)$$

where $\varphi^{(1)}_{i,k} = \left(\sum_{j=1}^{n} (\gamma_{ji,k})^+ \right)^2 + \left(\sum_{j=1}^{n} (\gamma_{ji,k})^- \right)^2$ *and* $\varphi^{(2)}_{i,k} = \left(\sum_{j=1}^{n} (\gamma_{ji,k})^+ \right) \left(\sum_{j=1}^{n} (\gamma_{ji,k})^- \right)$*, where* $(x)^+$ *and* $(x)^-$ *are the positive and negative parts of* x*, respectively.*

Proof See Appendix A.

Intuitively if the *ith* and *jth* entries of a eigenvector have the same sign, then the corresponding nodes react in a similar way to a shock propagating through the network (see, e.g. Estrada and Hatano 2008). Thus, the nodes can be partitioned into groups following the sign of the node-specific contribution $(\varphi_{i,k}^{(1)} - 2\varphi_{i,k}^{(2)})$ to the average centrality. At this point, we have completely developed the apparatus necessary to our state identification strategy based on network statistics. Regime identification is based on the restriction

$$\overline{q}(\tilde{G}_1) < \ldots < \overline{q}(\tilde{G}_K),$$

This constraint directly "separates" regimes according to the connectivity features of the network, which allows us to give a clear economic interpretation to the regimes: the first regime is associated with the lowest level of systemic risk and hence the lowest average of centrality scores across firms, the second regime corresponds to the next lowest average incidence of systemic risk, and so on, with the last regime associated with the strongest incidence of systemic risk.

4 Empirical Analysis

Our application focuses on all the constituents stocks of the S&P100 index for which we have at least fifteen years of continuous trading days as of the end of our sample, leaving us with $n = 83$ firms. The sample period is May 1996 - October 2014. The S&P 100 represents about 63% of the market capitalization of the S&P 500 and about half of the total market capitalization of the U.S. equity markets as of January 2017. These stocks tend to be the largest and most liquid companies in the U.S.

We analyze three popular linear asset pricing models starting from the plain vanilla CAPM, then extended to the three-factor model proposed by Fama and French (1993) to include both size and value factor-mimicking portfolios, to conclude with an implementation of Merton's (1973) intertemporal CAPM in which the aggregate dividend

yield, fixed income default and term spread are added to the excess returns on the market portfolio as priced factors. The default spread is computed as the difference between the yields of long-term corporate Baa bonds and long-term government bonds, and should reflect a risk premium for the aggregate risk of a firm's default on its debt. The term spread is measured as the difference between the yields on 10- and 1-year government bonds, and reflects the slope of the risk-free yield curve, a well-know business cycle leading indicator.[6]

4.1 Estimates of Latent States and Parameters

A priori, we assume that the latent states are persistent. This is based on the conventional wisdom that systemic risk is not a quickly mean-reverting process (see, e.g., Forbes and Rigobon 2002, Billio et al. 2012 and Diebold and Yilmaz 2014). In our application, we set the hyper-parameters to be rather uninformative: $m_k = \mathbf{0}$ and $M_k = 1000 I_n$ for each $k = 1, \ldots, K$. The prior for the hyper-inverse Wishart distribution is also set to be fairly vague, i.e., $c_k = 3$ and $C_k = 0.0001 I_n$. Finally, the marginal prior on the space of graphs is a Bernoulli distribution with $\psi = 2/(n-1)$ which provides a prior mode at n edges. To reduce sampling inefficiency we keep one in five draws and discard the initial 20% of the draws as burn-in sample. In a separate online Appendix we formally test for the number of regimes; the evidence from Bayes factors is clearly in favor of a model specification with two regimes.

Figures 1 shows the filtered probabilities of being in a state of high systemic risk across different linear factor specifications. The results indicate that the high connectivity regime started around 1998 and has strongly characterized the period 2000-2001 (i.e., the dot.com bubble, the market panic that followed the September 2001 terror attacks, the Enron and Worldcom financial scandals, and the unfolding of the events leading to the second Iraqi war), and the great financial crisis of 2008-2009.

[6]Data on corporate bonds and Treasuries are from the Fred II database at the Federal Reserve Bank of St.Louis. Following Campbell (1996), the innovations from a recursively estimated first-order Vector Auto-Regressive VAR(1) model are used as pricing factors in the Intertemporal CAPM implementation. As in Petkova (2006), these innovations are orthogonalized with respect to the excess return on the market portfolio and re-scaled such that they have the same variance.

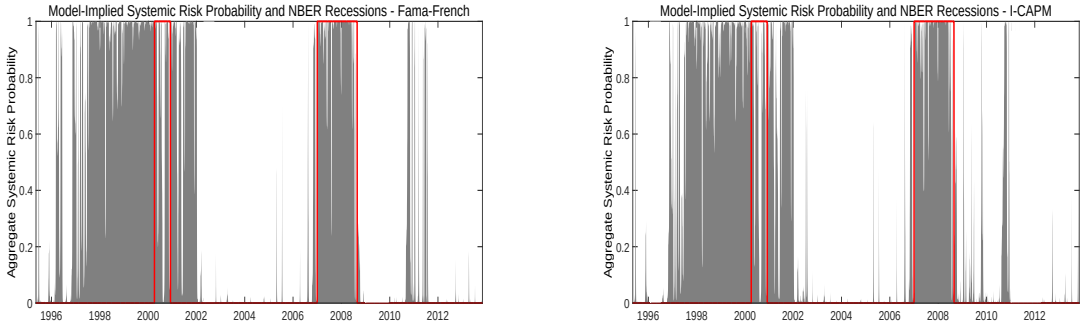


Figure 1. Filtered Probabilities of the High Systemic Risk Regime

The shaded areas represent the systemic risk filtered probabilities for a given model specification, while the red line shows the NBER recession indicator from the period following the peak through the trough. A value of 1 is a recessionary period, while a value of 0 is an expansionary period.

Although there is an obvious mis-match between the ex-post identification of the high network connectivity regime and the NBER business cycle indicator over the period 1998-2002, the NBER recession and high systemic risk tend to overlap during the recent financial crisis. As a whole, our filtered probabilities line up consistently with well-known periods of increasing turmoil in financial markets.[7]

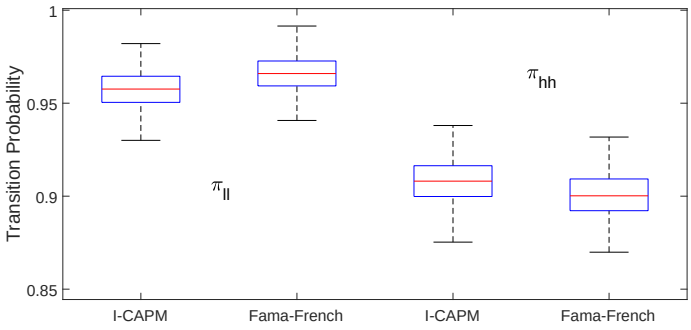


Figure 2. Transition Probabilities of the Latent States for Different Models

The first (last) two columns represent the posterior probability of staying in a state of low (high) systemic risk. The red line shows the posterior median estimates, the blue boxes mark horizontally the $25th$ and $75th$ quantiles.

[7]In this paper we are focusing on the fragility of the economic system rather than on financial crises per se. The fact that few events at the beginning of 2000 may have been reflected in a switch to a regime of high network connectivity is highly possible even though these events have not always triggered a full-blown financial crisis. Interestingly, we show below that also outside well known crisis episodes, financial firms (and a few other sectors) always occupy a central position in the network because of their high connectivity.

Figure 2 shows standard confidence box plots for the posterior transition probabilities of the latent systemic risk states across different model specifications. In both factor specifications, the states of high systemic risk are rather persistent, with an approximate median probability of $\pi_{hh} = 0.9$. The persistence of the high systemic risk regime is slightly lower than for the low systemic risk regime, which is compatible with the empirical evidence that turbulent periods of crisis and contagion tend to be less protracted than "normal" times.

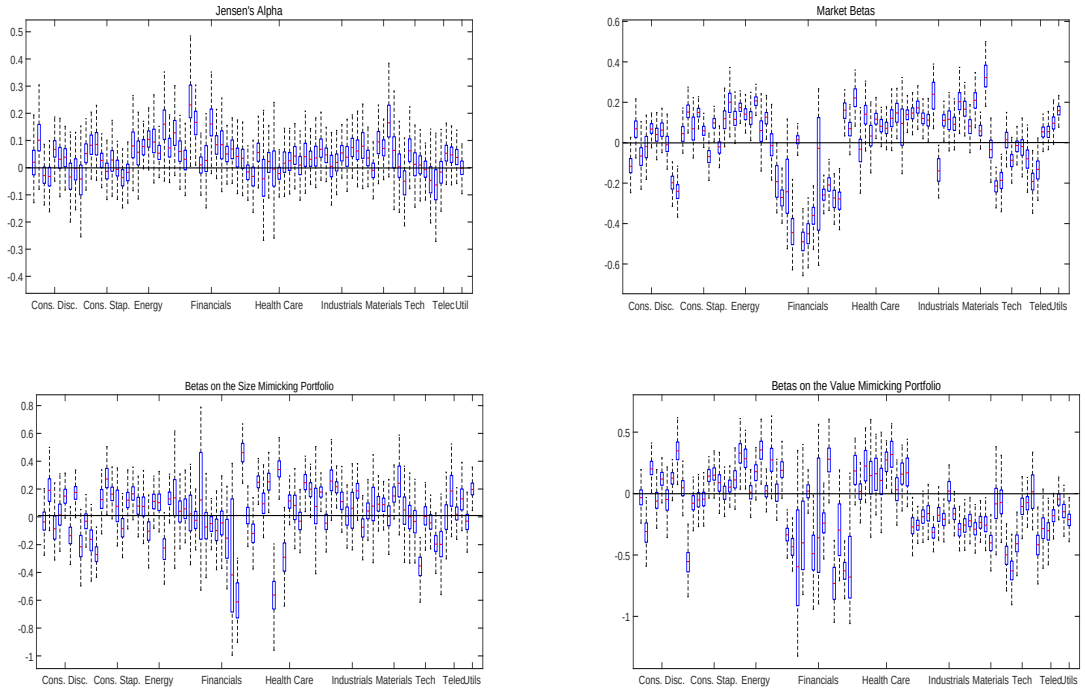


Figure 3. Changes in Betas Across Systemic Risk Regimes, Three-Factor Model

This figure shows the posterior of the between-state differences in the intercepts and regression betas. The top left (right) panel shows the posterior of the changes in the conditional intercepts (betas on the market portfolio). The bottom left and right panels report the changes in the betas for the size and value mimicking portfolios, respectively. We cluster the ticks on the horizontal axis according to an industry classification.

Figure 3 reports the box-plot of the posterior of the between-state differences, i.e. low minus high, in the conditional intercepts and in the regression betas from the Fama-French three-factor model. For ease of exposition and because they are similar to the CAPM, the results for the I-CAPM are reported in an online Appendix. Notice that, under the assumption of tradability of the factors, the conditional intercept

measures the unexpected excess stock returns, i.e. Jensen's alpha, while the betas measure the regime-dependent exposure to systematic risks. For each stock, we show the posterior distributions of the parameters as box plots and cluster these around industry-specific ticks on the horizontal axis.

The top-left panel shows that the Jensen's alphas tend to be higher when aggregate network connectivity is low. As a matter of fact, the differences between the regime-specific intercepts are predominantly positive, meaning that the inclusion of a sizable systemic risk component significantly increases the accuracy of standard factor pricing models (see, e.g., Ahern 2015). Moreover, the posterior estimates of the market betas (top-right panel) increase quite dramatically in regimes of high network connectivity. This tends to be more visible for companies in the Financial industry where the difference in the posterior mean of the market betas is as high as 0.3. A similar increase in the exposure to market risk is also estimated for some stocks in the Materials and Technology sectors. As far as the two remaining risk exposures are concerned, the Financial sector turns out to be more exposed to the value risk factor (bottom-right panel) when aggregate connectivity is high. The Industrial and Materials sectors also show an increasing exposure to the value mimicking portfolio in the high systemic risk regime.

4.2 Network Estimates

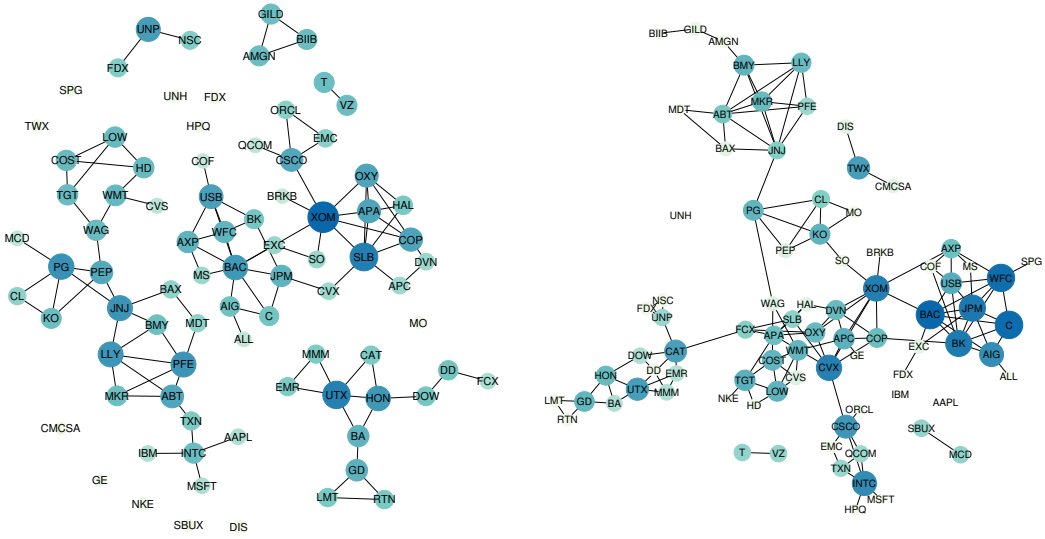
Under the MCMC estimation scheme outlined in Section 3, it is possible to define the posterior distribution of the graph and covariances $p(\Sigma_k, G_k|\mathbf{y}_{1:T})$ and to assess the statistical properties of the firm-specific network contribution using equation (21). Let $G_k^{(j)}$, $D_k^{(j)}$ and $\Sigma_k^{(j)}$ be the jth of J posterior draws of, respectively, the state-dependent graph, edge set and covariance matrix. As a result, it is possible to define the weight matrix $\tilde{\Sigma}_k^{(j)}$ following (19) and the corresponding weighted graph $\tilde{G}_k^{(j)} = (V, D_k^{(j)}, \tilde{\Sigma}_k^{(j)})$, such that the posterior centrality measure can be approximated as follows

$$\mathbb{E}(\overline{q}(G_k)|\mathbf{y}_{1:T}) = \int\int_{\mathcal{M}(G_k)\times\mathcal{G}} \overline{q}(\tilde{G}_k) p(\Sigma_k, G_k|\mathbf{y}_{1:T}) d\Sigma_k dG_k \approx \frac{1}{J}\sum_{j=1}^{J} \overline{q}(\tilde{G}_k^{(j)}) \qquad (25)$$

We first report the maximum a posteriori (MAP) estimates of the graphs. Figure 4 shows the results obtained for the Fama-French three-factor model. Financial firms become increasingly pivotal to the network during periods of high systemic risk. This is consistent with the conventional wisdom that the Financial sector is central to the transmission mechanism of shocks during a crisis. Instead, firms within the Energy sector show the highest degree of network centrality in regimes of low connectivity. The role of energy firms is not entirely unexpected, given that historically, the US economy has been exposed to quite a few energy or oil price shocks (see, e.g. Hamilton 2009 and Kilian 2009).[8] The results reported so far confirm that during periods of market turmoil, the systemic importance of the Financial sector substantially increases. Finally, the residual nature of the network identification strategy implicitly confirms that the importance of each firm holds after conditioning on size and value factors as sources of systematic risk. In other words, the key role of the Financial (Energy) sector when systemic risk is high (low) is confirmed even when size and value exposures are netted out.

We now shift our attention to the contribution of individual stocks/firms to aggregate systemic risk. Figure 5 shows the top 20 stocks ranked according to the posterior median of the weighted eigenvector centrality measure in (20). The left panel reports the results obtained from the Fama-French three-factor model. In a state of low aggregate network connectivity, energy stocks such as Exxon Mobil (XOM) and Schlumberger (SLB), tend to carry the highest weight in the system. Financials stocks rank in top positions when aggregate connectedness is high. For instance, Bank of America (BAC) has double the effect of Exxon Mobil (XOM) and almost four times the weight of Chevron (CVX) on system risk. The right panel of Figure 5 extends these results to the I-CAPM model. First, by conditioning on macroeconomic factors, the relative contribution of energy companies declines. Companies such as Anadarko Ptl. (APC), ConocoPhillips (COP), Occidental Ptl. (OXY), Apache (APA), and Schlumberger (SLB) show now a much lower centrality in the

[8]In an online Appendix we report the results for the I-CAPM implementation. The network tends to be more sparse for the I-CAPM where more regressors are included. This is possibly due to the residual nature of the estimated network. In this respect, by including a significant risk factor, the ability of our model to clean the network structure from spurious linkages increases.



(a) Low Systemic Risk (b) High Systemic Risk

Figure 4. Posterior Estimates of Network Connectivity, Three-Factor Model

This figures reports the maximum a posteriori (MAP) estimates of the graphs. The left (right) panel shows the structure of the network in the regime of low (high) cross-firm connectedness. The size and the color of the nodes are proportional to their weighted eigenvector centrality estimated as in (25). The darker (bigger) the color (size) of the node, the higher its contribution to aggregate systemic risk.

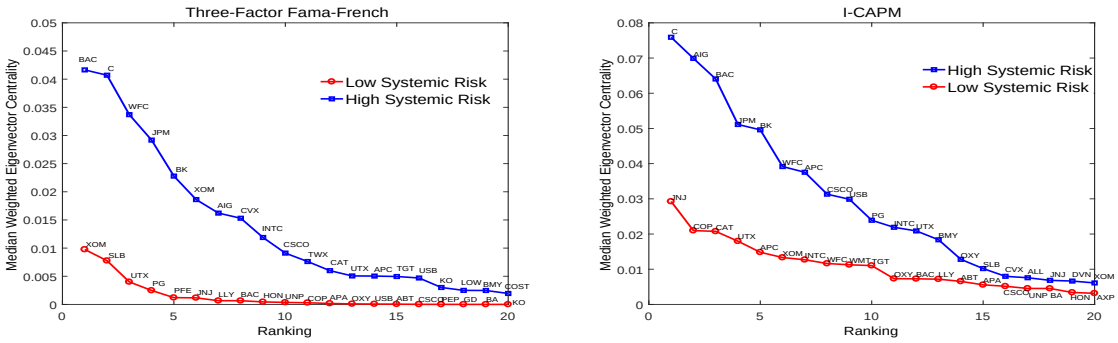


Figure 5. Firm-Level Network Centrality

This figure ranks the top 20 stocks according to the posterior median of the weighted eigenvector centrality measure in (20). The red (blue) line with circles (squares) shows firm-specific centrality measures when aggregate network connectedness is low (high).

posterior network. This suggests that traded factors in the three factor model are not sufficient to capture the aggregate state of the economy, which in turn appears to be inherently related to energy shocks. However, as in the Fama-French model, the weight of financial institutions exceeds all other industries' weights when systemic risk is high.

In a separate online Appendix, we report the results generalized at the industry level by taking the average over the posterior median over firm-specific weighted eigenvector centrality within each industry. We classify stocks in 11 sectors according to the Global Industry Classification Standard (GICS). The results confirm a substantial concentration of the network around few firms that belong to either the Financial and the Energy industries. The importance of the Utilities, Telecommunications, Health Care, Consumer Staples, and Consumer Discretionary sectors is almost negligible instead. Interestingly, the ranking of industries in terms of their systemic importance is robust across factor model specifications. A possible explanation is based on the dominance of the excess return on the market as priced risk factor, which is included in all of our linear factor specifications. In this respect, while additional factors may help to refine the analysis of the graph network, the key features of our empirical analysis depend on the fact that a market portfolio is included in the set of risk factors.

4.3 Market Value, Financial Losses and Network Centrality

The network centrality of a firm/industry is potentially linked to its relative market value. For instance, the relative equity weight of the Financial sector has dropped from 20% in 2006 to less than 10% by the end of the financial crisis of 2008-2009, when network connectivity has been high (see Figure 1). This may imply the existence of an inverse relationship between network centrality of the Financial sector and its market value. The opposite is true for the Energy sector: the relative market value of the Energy sector has increased throughout our sample and has tended to be higher during periods of high aggregate systemic risk. The same positive relationship applies, although to a weaker extent, to the Telecommunication industry, while

listed companies belonging to Industrial and Materials sectors do not display a clear mapping between valuations and systemic risk.

Because these empirical findings are rather stark, we proceed to test for the existence of a significant relationship between firm-level network centrality and market values across regimes. We do so by estimating a set of univariate cross-sectional regressions in which the explained and explanatory variables are the centrality measures for each firm and their average market value, respectively. We estimate these regressions for each factor pricing model and for each identified regime of network connectivity. We control for industry heterogeneity by including regime-specific industry fixed effects. Together with the cross-sectional regressions we compute the rank-correlation coefficient ρ as in Kendall (1938), which measures the correspondence between the ranking of the stocks based on their posterior centrality and their average market value in each regime. Table 1 reports the results.

Table 1. Network Centrality and Market Values

	CAPM				Fama-French				I-CAPM			
	Coeff	t-stat	R^2	ρ	Coeff	t-stat	R^2	ρ	Coeff	t-stat	R^2	ρ
High	0.021	1.201	0.012	0.054	0.011	1.291	0.021	0.061	0.021	1.254	0.013	0.045
Low	0.035	1.581	0.029	0.084	0.052	1.271	0.013	0.085	0.039	1.432	0.024	0.055

This table reports the regression estimates and the rank-correlation coefficient as in Kendall (1938). Standard errors are corrected for heteroskedasticity and autocorrelation in the residuals using a Newey-West HAC correction. Rank correlations ρ are highlighted in grey when the null hypothesis is rejected in tests with a size of 5% or lower.

We find evidence that systemic risk and market value are not precisely correlated. The slope coefficient is low in magnitude and statistically insignificant across regimes. The t-statistics are below the 5% significance level, and the adjusted R^2 is below 3% across models and regimes. One may argue that this finding does not square well with other evidence in the empirical finance literature. Indeed, the empirical evidence tends to find a positive association between the contribution to systemic risk and market capitalization especially for the banking and insurance sectors.

Although it is evident that the largest institutions caused major systemic effects during the 2008-2009 financial crisis, there have been other, earlier crises triggered by

smaller companies and sectors, e.g., dotcom crisis, which nonetheless increased the overall fragility of the economic system. In this respect, our paper is more about the estimation of systemic weakness through time-varying network connectivity rather than financial crises per se. Also, it should be clear that our goal here is not to over-throw a result from the empirical finance literature, but to deal with a potential concern that our empirical analysis may just represent a convoluted way to deal with firm and/or sector size and market values. Our results in Table 1 show that this is not the case: there is only weak correlation between market values and network centrality, although such correlation is indeed positive.

One additional feature that we expect from systemic risk measures is the ability to accurately explain the potential losses experienced by firms. To this end, we test the null hypothesis that firms more exposed to systemic risk are those that tend to record higher losses after major downturns. For each model specification, we regress the average maximum percentage financial loss on firm-specific centrality.[9] As above, we control for industry heterogeneity by including regime-specific industry fixed effects. The results are reported in Table 2.

Table 2. Network Centrality and Realized Financial Losses

	CAPM				Fama-French				I-CAPM			
	Coeff	t-stat	R^2	ρ	Coeff	t-stat	R^2	ρ	Coeff	t-stat	R^2	ρ
High	0.751	2.261	0.121	0.211	0.871	2.314	0.101	0.205	0.340	2.181	0.112	0.198
Low	0.412	1.901	0.145	0.181	0.301	1.913	0.062	0.171	0.456	1.859	0.061	0.169

This table reports the regression estimates and the rank-correlation coefficient as in Kendall (1938). Standard errors are corrected for heteroskedasticity and autocorrelation in the residuals using a Newey-West HAC correction. Rank correlations ρ are highlighted in grey when the null hypothesis is rejected in tests with a size of 5% or lower.

We find that companies more exposed to the overall systemic risk more likely suffer significant losses when aggregate systemic risk is high. The cross-sectional regression coefficient is indeed significant at standard size levels and the adjusted R^2 is

[9]The maximum percentage loss for a firm is defined to be the maximum difference between the market capitalization of an institution between time t and $t+h$ dividend by its market capitalization at time t. Such an average measure is computed by averaging out the maximum percentage loss for a given regime k of systemic risk.

around 12% across models, which means that firms that are more contemporaneously interconnected with the rest of the market are also those that experience major losses in periods of high systemic risk. In this respect, our centrality measure is similar to the marginal expected shortfall measure originally proposed by Acharya, Pedersen, Phillippon, and Richardson (2017), which tracks the sensitivity of returns of stock i to a system-wide extreme event, thereby providing a market-based measure of fragility exposure of a firm. However, such positive correlation between network centrality and market losses is less significant when aggregate connectedness decreases, i.e., in states of low systemic risk.

For each regime, Table 2 also reports the Kendall's rank-correlation coefficients ρ between the ranking of firms according to centrality and losses. The results confirm that there is a significant relationship between network centrality and value losses across firms, especially during periods of high aggregate systemic risk. This is consistent with previous evidence in Billio et al. (2012) and Diebold and Yilmaz (2014), the theoretical framework of Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012), and the conventional wisdom that firms more exposed to systemic risks are bound to face larger losses on average.

5 Conclusions

In the aftermath of the great financial crisis, one of the main questions for economists and market participants has concerned the extent to which the economy is robust to unexpected shocks. In the language of network analysis, this translates into a desire to understand the nature and patterns of cross-firm connectivity. We address this question by developing a novel Markov Switching Graphical Seemingly Unrelated Regression (MS-GSUR) model, which allows us to jointly estimate standard SUR-type relationships of firms' asset price connectedness from the error terms of a range of popular linear multi-factor pricing model specifications. By conditioning on different sources of systematic risk, we implicitly recognize that systematic and systemic risk might be conditionally independent but not mutually exclusive.

Methodologically, we develop a Markov Chain Monte Carlo (MCMC) scheme which allows to sample the posterior estimates of all parameters and measures of interest. The label-switching identification problem is solved using the graph-theoretic properties of the state-specific conditional dependence structure of the regression residuals. In this respect, we propose a new weighted eigenvector centrality measure, which accounts not only for the number of adjacent nodes, but also for the weights of the edges and for the number of indirect connections between nodes. More generally, our new measure implies that the existence of a financial linkage between firms is encoded in the presence or absence of an edge, while the strength of the linkage is measured by their corresponding covariance terms.

In the empirical analysis we show that the dynamics of systemic risk for the constituents of the S&P100 index over the period 1996-2014 can be captured by two regimes, in which a state of high connectedness characterizes the period 1999-2003 and the great financial crisis of 2008-2009. In addition, we show that a few financial institutions heavily outweighed other firms in the network during these periods and that shocks to the Financial sector turned out to be the most systemically important. Finally, both cross-sectional regressions and a rank-correlation analysis show that, while market capitalization does not significantly explain centrality, those firms more exposed to the overall systemic risk more likely suffer significant losses when aggregate systemic risk is high.

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012). The network origins of aggregate fluctuations. *Econometrica 80*(5), 1977–2016.

Acharya, V., L. Pedersen, T. Phillippon, and M. Richardson (2017). Measuring systemic risk. *Review of Financial Studies 30*(1), 2–47.

Ahelegbey, D. F., M. Billio, and R. Casarin (2016). Bayesian graphical models for structural vector autoregressive processes. *Journal of Applied Econometrics 31*(2), 357–386.

Ahern, K. (2015). Network centrality and the cross-section of stock returns. *Working Paper*.

Barigozzi, M. and C. Brownlees (2014). Nets: Network estimation for time series. *Unpublished Working Paper*.

Barrat, A., M. Barthlemy, R. Pastor-Satorras, and A. Vespignani (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America 101*(11), 3747–3752.

Billio, M., M. Getmansky, A. W. Lo, and L. Pelizzon (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics 104*(3), 535–559.

Bollobás, B. (1998). *Modern Graph Theory*. Springer.

Bollobás, B. (2001). *Random Graphs*. Cambridge University Press.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology 2*(1), 113–120.

Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social networks 29*(4), 555–564.

Brownlees, C., E. Nualart, and Y. Sun (2014). Realized networks. *Unpublished Working Paper*.

Campbell, J. Y. (1996). Understanding risk and return. *Journal of Political economy 104*(2), 298–345.

Carter, C. and R. Kohn (1994). On gibbs sampling for state-space models. *Biometrika 81*(3), 541–553.

Carvalho, C., M. West, and H. Massam (2007). Simulation of hyper-inverse wishart distributions in graphical models. *Biometrika 94*(3), 647–659.

Carvalho, C. M., M. West, et al. (2007). Dynamic matrix-variate graphical models. *Bayesian analysis 2*(1), 69–97.

Casella, G. and C. P. Robert (2004). *Monte Carlo Statistical Methods*. New York: Springer Verlag.

Chib, S. (1995). Marginal likelihood from the gibbs output. *Journal of the American Statistical Association 90*(432), 1313–1321.

Cochrane, J. H. (2009). *Asset Pricing:(Revised Edition)*. Princeton university press.

Corsetti, G., M. Pericoli, and M. Sbracia (2005). Some Contagion, some Interdependence: More Pitfalls in Tests of Financial Contagion. *Journal of International Money and Finance 24*(8), 1177 – 1199.

Corsetti, G., M. Pericoli, and M. Sbracia (2011). *Financial Contagion: The Viral Threat to the Wealth of Nations*, Chapter Correlation Analysis of Financial Contagion, pp. 11–20. John Wiley & Sons.

Dawid, A. and S. Lauritzen (1993). Hyper-markov laws in the statistical analysis of decomposable graphical models. *The Annals of Statistics 21*(3), 1272–1317.

Dempster, A. (1972). Covariance selection. *Biometrics 28*(1), 157–175.

Diebold, F. and K. Yilmaz (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics 182*(1), 119–134.

Diebold, F. and K. Yilmaz (2015). *Financial and Macroeconomic Connectedness: A Network Approach to Measurement and Monitoring*. Oxford University Press.

Estrada, E. and N. Hatano (2008). Communicability in complex networks. *Physical Review E 77*(3), 036111–12.

Estrada, E. and N. Hatano (2009). Communicability graph and community structures in complex networks. *Applied Mathematics and Computation 214*(2), 500–511.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics 33*(1), 3–56.

Forbes, K. and R. Rigobon (2000). *Measuring Contagion: Conceptual and Empirical Issues* (International Financial Contagion ed.). Kluwer Academic Publisher.

Forbes, K. and R. Rigobon (2002). No contagion, only interdependence: Measuring stock market comovements. *Journal of Finance 57*(5), 2223–2261.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports 486*(35), 75 – 174.

Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis 15*(2), 183–202.

Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. Berlin: Springer-Verlag.

Giudici, P., Green, and PJ (1999). Decomposable graphical gaussian model determination. *Biometrika 86*(4), 785–801.

Guidolin, M. and A. Timmermann (2008). International asset allocation under regime switching, skew and kurtosis preferences. *Review of Financial Studies 21*(2), 889–935.

Hamilton, J. (1994). *Time series analysis*. Princeton University Press.

Hamilton, J. (2009). Understanding crude oil prices. *The Energy Journal*, 179–206.

Hammersley, J. and P. Clifford (1971). Markov fields on finite graphs and lattices. *Unpublished Manuscript, Oxford University*.

Han, L., F. Escolano, E. R. Hancock, and R. C. Wilson (2012). Graph characterizations from Von Neumann entropy. *Pattern Recognition Letters 33*(15), 1958 – 1967.

Hautsch, N., J. Schaumburg, and M. Schienle (2015). Financial network systemic risk contributions. *Review of Finance 19*(2), 685–738.

Jones, B., C. Carvalho, A. Dobra, C. Hans, C. Carter, and M. West (2005). Experiments in stochastic computation for high-dimensional graphical models. *Statistical Science 20*(4), 388–400.

Jones, B. and M. West (2005). Covariance decomposition in undirected Gaussian graphical models. *Biometrika 92*(4), 779–786.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika 30*(1/2), 81–93.

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *The American economic review 99*(3), 1053–1069.

Lauritzen, S. (1996). *Graphical Models*. Claredon Press, Oxford.

Petkova, R. (2006). Do the fama-french factors proxy for innovations in predictive variables? *Journal of Finance 61*(2), 581–612.

Qi, X., E. Fuller, R. Luo, and C.-Q. Zhang (2015). A novel centrality method for weighted networks based on the kirchhoff polynomial. *Pattern Recognition Letters 58*(1), 51–60.

Qi, X., E. Fuller, Q. Wu, Y. Wu, and C.-Q. Zhang (2012). Laplacian centrality: A new centrality measure for weighted networks. *Information Sciences 194*(1), 240 – 253.

Roberts, G. and S. Sahu (1997). Updating schemes, covariance structure, blocking and parametrisation for the gibbs sampler. *Journal of the Royal Statistical Society 59*(2), 291 – 318.

Rodriguez, A., A. Lenkoski, and A. Dobra (2011). Sparse covariance estimation in heterogeneous samples. *Electronic Journal of Statistics 5*, 981.

Rubinov, M. and O. Sporns (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage 52*(3), 1059 – 1069.

Timmermann, A., D. Blake, I. Tonks, and A. Rossi (2014). Network centrality and fund performance. *CFR Working Papers N 15-16, University of Cologne, Centre for Financial Research*.

Wang, H. (2010). Sparse seemingly unrelated regression modelling: Applications in finance and econometrics. *Computational Statistics and Data Analysis 54*(11), 2866–2877.

Wang, H., C. Reeson, and C. M. Carvalho (2011). Dynamic financial index models: Modeling conditional dependencies via graphs. *Bayesian Analysis 6*(4), 639–664.

Wang, H. and M. West (2009). Bayesian analysis of matrix normal graphical models. *Biometrika 96*(4), 821–834.

Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics*. Chichester, UK: John Wiley and Sons.

Appendix

A Proofs of the results of Section 3

A.1 Proof of the result in 14

The complete likelihood of the data is then defined as

$$p\left(\mathbf{y}_{1:T}, \mathbf{s}_{1:T}|\boldsymbol{\theta}, G\right) = \prod_{t=1}^{T} p\left(\mathbf{y}_t|s_t, \boldsymbol{\theta}, G\right) p\left(s_t|s_{t-1}, \boldsymbol{\theta}\right) p\left(s_0\right) \tag{A.26}$$
$$= \prod_{t=1}^{T} \left(2\pi\right)^{-n/2} \left|\Sigma(s_t)\right|^{-1/2} \exp\left(-\frac{1}{2}\left(\mathbf{y}_t - Z_t'\boldsymbol{\beta}(s_t)\right)' \Sigma_t^{-1}\left(\mathbf{y}_t - Z_t'\boldsymbol{\beta}(s_t)\right)\right) \prod_{k,l=1}^{K} \pi_{lk}^{\xi_{lk,t}} \pi_0,$$

with $\xi_{lk,t} = \mathbb{I}_{\{k\}}\left(s_t\right)\mathbb{I}_{\{l\}}\left(s_{t-1}\right)$. Thus the full conditional of Σ_k given $\mathbf{y}_{1:T}$, $\boldsymbol{\theta}$, $\mathbf{s}_{1:T}$, is

$$p\left(\Sigma_k|\mathbf{y}_{1:T}, \boldsymbol{\theta}, \mathbf{s}_{1:T}, \boldsymbol{\beta}_k\right) \propto \prod_{t=1}^{T} \left(2\pi\right)^{-n/2} \left|\Sigma(s_t)\right|^{-1/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(\Sigma(s_t)^{-1}\mathbf{e}_t\mathbf{e}_t'\right)\right\} p(\Sigma_k)$$
$$\propto \prod_{t\in\mathcal{T}_k} \left|\Sigma_k\right|^{-1/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(\Sigma_k^{-1}\mathbf{e}_{tk}\mathbf{e}_{tk}'\right)\right\} p(\Sigma_k) \tag{A.27}$$

where $\mathcal{T}_k = \{t : s_t = k\}$, $\mathbf{e}_t = \mathbf{y}_t - Z_t'\boldsymbol{\beta}(s_t)$ and $\mathbf{e}_{tk} = \mathbf{y}_t - Z_t'\boldsymbol{\beta}_k$. Exploiting the conditional independence structure encoded in k-th state graph G_k it follows

$$\begin{aligned}
&p\left(\Sigma_k|\mathbf{y}_{1:T}, \boldsymbol{\theta}, \mathbf{s}_{1:T}, \boldsymbol{\beta}_k\right) \propto \\
&\propto \prod_{j=1}^{n_{A,k}} \left|\Sigma_{A_{j,k}}\right|^{-T_k/2} \exp\left\{-\frac{1}{2}\Sigma_{A_{j,k}}^{-1} E_{A_{j,k}}^*\right\} \times \prod_{j=1}^{n_{A,k}} \left|\Sigma_{A_{j,k}}\right|^{-\left(c_k+2T_{A_{j,k}}\right)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma_{A_{j,k}}^{-1} C_{A_{j,k}})\right\} \cdot \\
&\quad \cdot \prod_{j=1}^{n_{B,k}} \left(\left|\Sigma_{B_{j,k}}\right|^{-\left(c_k+2T_{B_{j,k}}\right)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma_{B_{j,k}}^{-1} C_{B_{j,k}})\right\}\right)^{-1} \\
&\propto \prod_{j=1}^{n_{A,k}} \left|\Sigma_{A_{j,k}}\right|^{-\left(c_k+2T_{A_{j,k}}+T_k\right)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma_{A_{j,k}}^{-1}(C_{A_{j,k}} + E_{A_{j,k}}^*)\right\} \cdot \\
&\quad \cdot \prod_{j=1}^{n_{B,k}} \left(\left|\Sigma_{B_{j,k}}\right|^{-\left(c_k+2T_{B_{j,k}}\right)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma_{B_{j,k}}^{-1} C_{B_{j,k}})\right\}\right)^{-1} \\
&\propto \mathcal{HIW}_{G_k}\left(c_k + T_k, C_k + E_k^*\right),
\end{aligned} \tag{A.28}$$

where $E_{A_{j,k}}^*$ is the block of E_k^* corresponding to $\Sigma_{A_{j,k}}$, $T_{B_{j,k}} = Card(B_{j,k})$ and $\Sigma_{B_{j,k}}$ is defined over the set of separators as (9).

A.2 Proof of Proposition 1

The proof of Proposition 1 is composed of two parts. First we show that our weighted eigenvector centrality measure is the limit of the rescaled Bonacich's $c(b)$ centrality (see Bonacich 2007) of the weighted graph $\tilde{G}_k$ of regime $k \in K$. In the second part, we show that the Bonacich's $c(b)$ centrality can be written as a weighted sum over all walks between all pairs of nodes of the graph $\tilde{G}_k$. Let $\tilde{\sigma}_{ij,k}$ be the i-th row j-th column entry of $\tilde{\Sigma}_k$. The relative weighted centrality scores of $\tilde{G}_k$ can defined by

$$\lambda\gamma_i = \sum_{j=1}^{n} \tilde{\sigma}_{ij,k}\gamma_j, \quad i = 1, \ldots, n$$

which can be re-written in a more compact format as the eigenvector equation $\tilde{\Sigma}_k\boldsymbol{\gamma} = \lambda\boldsymbol{\gamma}$ with n solutions given by the eigenvalues $\lambda_{i,k}$, $i = 1, \ldots, n$ with $\lambda_{1,k} \leq \ldots \leq \lambda_{n,k}$, and the associated eigenvectors $\boldsymbol{\gamma}_{ik} = (\gamma_{1i,k}, \ldots, \gamma_{ni,k})'$ $i = 1, \ldots, n$. Since $\tilde{\Sigma}_k$ is symmetric, the decomposition $\tilde{\Sigma}_k^l = \boldsymbol{\gamma}_k\Lambda_k^l\boldsymbol{\gamma}_k'$ holds with $\Lambda_k = \text{diag}\{\lambda_{1,k}, \ldots, \lambda_{n,k}\}$ and $\boldsymbol{\gamma}_k = (\boldsymbol{\gamma}_{1k}, \ldots, \boldsymbol{\gamma}_{nk})$, with $\boldsymbol{\gamma}_{i,k}$ $i = 1, \ldots, n$ being the orthonormal eigenvectors, $\gamma_{ij,k}$ the ith element of the jth eigenvector for the state$-k$ weighted covariance $\tilde{\Sigma}_k$.

The Bonacich's $c(b)$ centrality can be defined as

$$c(b) = \lim_{L\to\infty} \sum_{l=1}^{L} b^{l-1}\tilde{\Sigma}_k^l\boldsymbol{\iota},$$

with $\boldsymbol{\iota}$ the n-dimensional unit vector. By applying the decomposition of $\tilde{\Sigma}_k$ given above

$$\begin{aligned} c(b) &= \lim_{L\to\infty} \frac{1}{b} \sum_{l=1}^{L} b^l(\boldsymbol{\gamma}_k\Lambda_k^l\boldsymbol{\gamma}_k')\boldsymbol{\iota} = \lim_{L\to\infty} \frac{1}{b} \sum_{l=1}^{L} b^l \left(\sum_{i=1}^{n} \boldsymbol{\gamma}_{i,k}\lambda_{i,k}^l\boldsymbol{\gamma}_{i,k}'\right)\boldsymbol{\iota} = \\ &= \frac{1}{b} \sum_{i=1}^{n} \left(\lim_{L\to\infty} \sum_{l=1}^{L} b^l\lambda_{i,k}^l\right) \boldsymbol{\gamma}_{i,k}\boldsymbol{\gamma}_{i,k}'\boldsymbol{\iota}, \end{aligned} \tag{A.29}$$

Assuming orthonormal eigenvectors and $|b| < 1/\lambda_{n,k}$, it follows that (see, e.g. Bonacich 2007):

$$c(b) = \frac{1}{b} \sum_{i=1}^{n} \varphi_i(b,k)\boldsymbol{\gamma}_{i,k}, \quad \text{where} \quad \varphi_i(b,k) = \frac{b\lambda_{i,k}}{1 - b\lambda_{i,k}}$$

Define $I_\delta(1/\lambda_{n,k}) = \{b : 0 < 1/\lambda_{n,k} - b < \delta\}$, $\delta > 0$ as a neighborhood of the inverse maximal eigenvalue $\lambda_{n,k}$. Then $\forall\varepsilon > 0$, $\exists\delta_j > 0$ such that for $\forall b \in I_{\delta_j}(1/\lambda_{n,k})$, $|\varphi_j(b,k)/\kappa(b,k)| < \varepsilon/n$, $j \neq n$ and $\exists\delta_n > 0$ such that for $\forall b \in I_{\delta_n}(1/\lambda_{n,k})$, $|\varphi_n(b,k)/\kappa(b,k) - 1| < \varepsilon/n$, where $\kappa(b,k) = (\sum_{i=1}^{n} \varphi_i(b,k))$. Choosing $\delta = \min\{\delta_j, j = 1, \ldots, n\}$, and by the triangle inequality, it follows that

$$\left|\left|\frac{1}{\kappa(b,k)}c(b) - \boldsymbol{\gamma}_{n,k}\right|\right|_2 < \left|\frac{1}{\kappa(b,k)}\right| \left(\sum_{i=1}^{n-1} |\varphi_i(b,k)| + |\frac{1}{\kappa(b,k)}\varphi_n(b,k) - 1|\right) < \varepsilon, \tag{A.30}$$

where $\|\mathbf{x}\|_2$ denotes the L^2-norm of $\mathbf{x}$. To conclude the first part of the proof we need to find the relationship between our centrality measure $\overline{q}(\tilde{G}_k)$ and the limit given above. From the definition

of $\overline{q}(\tilde{G}_k)$

$$\overline{q}(\tilde{G}_k) = \frac{1}{n}\boldsymbol{\iota}'\boldsymbol{\gamma}_{n,k} \tag{A.31}$$

and by Cauchy-Schwarz inequality

$$\left|\frac{1}{n\kappa(b,k)}\boldsymbol{\iota}'c(b) - \overline{q}(\tilde{G}_k)\right| < \left|\left|\frac{1}{n}\boldsymbol{\iota}\right|\right|_2 \left|\left|\sum_{i=1}^{n}\left(\frac{1}{\kappa(b,k)}\varphi_i(b,k)\boldsymbol{\gamma}_{i,k}\right) - \boldsymbol{\gamma}_{n,k}\right|\right|_2 \tag{A.32}$$

we conclude that $\forall\varepsilon > 0$, $\exists\delta > 0$ such that for $\forall b \in I_\delta(1/\lambda_{n,k})$

$$\left|\frac{1}{n\kappa(b,k)}\boldsymbol{\iota}'c(b) - \overline{q}(\tilde{G}_k)\right| < \varepsilon \tag{A.33}$$

As for the second part of the proof, we can use Definitions 1 and 2 in the main text and rewrite part the right-hand side of (A.31) as:

$$\boldsymbol{\iota}'c(b) = \lim_{L\to\infty}\frac{1}{b}\sum_{l=1}^{L} b^l\boldsymbol{\iota}'\tilde{\Sigma}_k^l\boldsymbol{\iota} \tag{A.34}$$

which can be further re-written as a function of the sum of the weights over all possible walks between all pairs of nodes:

$$\boldsymbol{\iota}'\tilde{\Sigma}_k^l\boldsymbol{\iota} = \sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{\sigma}_{ij,k}^{(l)} = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{i_1\in V_k}\tilde{\sigma}_{ii_1,k}^{(l-1)}\tilde{\sigma}_{i_1j,k} \tag{A.35}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{i_1\in V_k}\sum_{i_2\in V_k}\tilde{\sigma}_{ii_2,k}^{(l-2)}\tilde{\sigma}_{i_2i_1,k}\tilde{\sigma}_{i_1j,k} = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{i_1\in V_k}\cdots\sum_{i_{l-1}\in V_k}\prod_{r=1}^{l}\tilde{\sigma}_{i_{r-1}i_r,k}\right) \tag{A.36}$$

Let $a_{i_{r-1}i_r,k}$ be the (i_{r-1}, i_r)-th element of the adjacency matrix A_k associated with G_k and $P_{ij,k}^{(l)} = \{(D_k(p), V_k(p));\, V_k(p) = \{i_0, \ldots, i_l\} \subset V_k,\, D_k(p) = \{e_1, \ldots, e_l\} \subset E, i_0 = i, i_l = j\}$ be the set of all walks of length l between i and j. Since $\prod_{r=1}^{l}\tilde{\sigma}_{i_{r-1}i_r,k} = \prod_{r=1}^{l} a_{i_{r-1}i_r,k}\sigma_{i_{r-1}i_r,k}$ and

$$\prod_{r=1}^{l} a_{i_{r-1}i_r,k} = \begin{cases} 1 & \text{if } (i_0, e_1, \ldots, e_l, i_l) \in P_{ij,k}^{(l)} \\ 0 & \text{otherwise} \end{cases}$$

it follows that

$$\boldsymbol{\iota}'\tilde{\Sigma}_k^l\boldsymbol{\iota} = \sum_{i,j\in V}\sum_{p\in P_{ij,k}^{(l)}}\tilde{\sigma}(p) \tag{A.37}$$

where $\tilde{\sigma}(p) = \prod_{r=1}^{l}\tilde{\sigma}_{i_{r-1}i_r,k}$ is the path weight. By plugging (A.37) in (A.34) and substituting the result in (A.31) we obtain (22) in the text.

A.3 Proof of Corollary 1

From Proposition 1, our centrality measure $\overline{q}(\tilde{G}_k)$ has the following limit

$$\frac{1}{\kappa(b,k)}\sum_{l=1}^{\infty} b^l \boldsymbol{\iota}'\tilde{\Sigma}_k^l \boldsymbol{\iota} = \frac{1}{\kappa(b,k)}\sum_{i,j\in V_k}\sum_{l=1}^{\infty} b^l \sum_{p\in P_{ij,k}^{(l)}} \tilde{\sigma}(p), \quad \text{where} \quad \tilde{\sigma}(p) = \prod_{r=1}^{l} \tilde{\sigma}_{i_{r-1}i_r,k} \tag{A.38}$$

as b tends to $1/\lambda_{n,k}$ from below. Then, since it is possible to obtain a path from a walk by removing its sub-cycles we can define for each pair of nodes i and j the set of paths $P_{ij,k}^{*}$ between i and j with minimum length l_{ij}^{*} (shortest paths) and the set $P_{ij,k}^{(l)}$ of all remaining walks between the two nodes with length $l > l_{ij}^{*}$. Since $P_{ij,k}^{(l)} = \emptyset$ for all $l < l^{*}$, it follows that

$$\sum_{l=1}^{\infty} b^l \sum_{p\in P_{ij,k}^{(l)}} \prod_{r=1}^{l} \tilde{\sigma}_{i_{r-1}i_r,k} = \left(\sum_{p\in P_{ij,k}^{*}} b^{l_{ij}^{*}-1} \prod_{r=1}^{l} \tilde{\sigma}_{i_{r-1}i_r,k} + \sum_{l>l_{ij}^{*}}^{\infty} b^{l-1} \sum_{p\in P_{ij,k}^{(l)}} \prod_{r=1}^{l} \tilde{\sigma}_{i_{r-1}i_r,k} \right)$$

which gives the desired decomposition.

A.4 Proof of Corollary 2

The limit of our weighted eigenvector centrality measure $\overline{q}(\tilde{G}_k)$, as b tends to $1/\lambda_{n,k}$ from below, can be re-written as

$$\frac{1}{\kappa(b,k)}\frac{1}{b}\sum_{i=1}^{n}\sum_{l=1}^{\infty}(b\lambda_{i,k})^l \boldsymbol{\iota}'\boldsymbol{\gamma}_{i,k}\boldsymbol{\gamma}_{i,k}'\boldsymbol{\iota} \tag{A.39}$$

Let us denote with $\boldsymbol{\gamma}_{i,k}^{+}$ the n-dimensional vector with j-th element $\gamma_{ji,k}^{+} = (\gamma_{ji,k})^{+}$, $j = 1,\ldots,n$ and with $\boldsymbol{\gamma}_{i,k}^{-}$ the n-dimensional vector with j-th element $\gamma_{ji,k}^{+} = (\gamma_{ji,k})^{+}$, $j = 1,\ldots,n$, where $(x)^{+}$ and $(x)^{-}$ denote the positive and negative part of x, respectively and $\gamma_{ji,k}$ the j-th element of $\boldsymbol{\gamma}_{i,k}$. Then, $\boldsymbol{\iota}'\boldsymbol{\gamma}_{i,k} = \boldsymbol{\iota}'(\boldsymbol{\gamma}_{i,k}^{+} - \boldsymbol{\gamma}_{i,k}^{-})$ and

$$\begin{aligned}
&\sum_{i=1}^{n}\frac{b\lambda_{i,k}}{1-b\lambda_{i,k}}\boldsymbol{\iota}'\boldsymbol{\gamma}_{i,k}\boldsymbol{\gamma}_{i,k}'\boldsymbol{\iota} = \sum_{i=1}^{n}\frac{b\lambda_{i,k}}{1-b\lambda_{i,k}}(\boldsymbol{\iota}'\gamma_{i,k}^{+})^2 + \sum_{i=1}^{n}\frac{b\lambda_{i,k}}{1-b\lambda_{i,k}}(\boldsymbol{\iota}'\boldsymbol{\gamma}_{i,k}^{-})^2 \\
&- 2\sum_{i=1}^{n}\frac{b\lambda_{i,k}}{1-b\lambda_{i,k}}\boldsymbol{\iota}'\gamma_{i,k}^{+}\boldsymbol{\iota}'\gamma_{i,k}^{-}
\end{aligned}$$

which yields the decomposition for $\overline{q}(\tilde{G}_k)$ given in Eq. (24) with $\varphi_{i,k}^{(1)} = (\boldsymbol{\iota}'\gamma_{i,k}^{+})^2 + (\boldsymbol{\iota}'\gamma_{i,k}^{-})^2 = (\sum_j^n(\gamma_{ji,k})^{+})^2 + (\sum_j^n(\gamma_{ji,k})^{-})^2$ and $\varphi_{i,k}^{(2)} = \boldsymbol{\iota}'\gamma_{i,k}^{+}\boldsymbol{\iota}'\gamma_{i,k}^{-} = (\sum_j^n(\gamma_{ji,k})^{+})(\sum_j^n(\gamma_{ji,k})^{-})$.